Do Happy Words Sound Happy? A study of the relation between form and meaning for English words expressing emotions

Vivi Nastase
EML Research gGmbH
Heidelberg, Germany
nastase@eml-research.de

Marina Sokolova
Université de Montreal
Montreal, Quebec, Canada
sokolovm@iro.umontreal.ca

Jelber Sayyad Shirabad
University of Ottawa
Ottawa, Ontario, Canada
jsayyad@site.uottawa.ca

Abstract

This paper presents a study of the relation between a word's form and the emotion it expresses. We analyze the possibility that the form of words expressing emotions is not completely arbitrary, but in fact, their sound evokes the emotion conveyed. We explore the relation between word form and emotions using a variety of word form representations and machine learning methods. We first show that words expressing an emotion are more similar among them than with words expressing other emotions, and then we discuss the sounds of emotions.

Keywords

sentiment analysis, word form, meaning, machine learning

1 Introduction

A word has two components: **word form**, a sequence of sounds (pronunciation) and, possibly, letters/characters (written form), and **meaning**. The word form is also called **signifier**, and its meaning, or referent in the world, is called **signified**: the word form *tree* with the pronunciation */trE/*[1] has as referent in the real world a TREE entity.[2] While it is usually accepted that the relation between signifier and signified is largely arbitrary [5], the idea that sounds may carry meaning has appeared at several points in time [8], and is still a matter of debate and research.

In this paper we study the relationship between signifier and signified for a class of words which can be particularly susceptible to the way a word sounds: words that express emotions – either positive or negative, or a more fine grained range (anger, disgust, fear, joy, sadness, surprise).

We work with data annotated with emotion tags: WordNet Affect and the dictionary from the Linguistic Inquiry and Word Count system. We work with the pronunciation and written form of a word. We represent the word form in various ways, using separately the written and pronunciation versions. We investigate the connection between form and emotion conveyed in two steps. We first verify, through machine learning experiments, whether such a connection exists. The results support this hypothesis, by showing that words expressing the same emotion have more in common with each other than with words expressing other emotions. In a second step, we analyze whether the sounds of happy words are indeed happy sounding. This is a harder question to answer, as perception is subjective. We discuss the sounds of emotions based on the most salient features in our experiments and research on emotion recognition in speech.

Apart from a purely theoretical benefit, finding a relation between the way the words sound and the emotion expressed contributes to research in sentiment analysis, very much part of the highly explored areas of NLP these days, authorship analysis and other research areas. From a practical point of view, such relations could be exploited in advertising, where product names that have no literal meaning rely on their sound to catch the attention and desire of potential customers [1].

2 Motivation

It is a long held belief that the association between a word-form and its meaning is arbitrary [5]: there is nothing about a TREE that evokes the sequence of letters or sounds that form the English word *tree*. Support of this theory comes from language variation: a TREE is called *tree* in English, but *Baum* in German, *albero* in Italian, and numerous other variants in the languages of the world. If there was anything intrinsic to TREE that would link it to the form *tree*, it would have been called the same in all languages.

There are also onomatopoeic words, which sound like the concept they describe [2]. Onomatopoeia are language specific. In English lions *roar*, cats *purr*, flies *buzz*, snakes *hiss*, fireworks go *boom* and *bang*.

In between the two extremes of total arbitrariness of form relative to meaning and identity of the two, there are *mellifluous* words. Coming from the Latin *mellifluus* = *mel*(honey)+*fluere*(flow) – dripping with honey – mellifluous has come to refer to words whose sounds evoke the concepts they refer to. Such words were particularly exploited for effects in poetry [17]. We also use them in our everyday speech: we *hush* to make silence, we *mumble* when we speak in a low inarticulate manner.

Arbitrariness of the connection between sound and meaning is not universally accepted. The theory of *sound symbolism* or *phonosemantics*, according to which most words in a language fall into a category similar to *mellifluous* – every sound carries a certain meaning, which evoke certain aspects of a concept whose name contains this sound – has ancient roots. Plato, through his characters in the Cratylus dialogue

[1] From the online version of the Merriam-Webster: http://www.m-w.com .
[2] For the remainder of the paper, the signifier will be written in *italics*, and the signified in SMALLCAPS.

– Hermogenes and Socrates – discusses the provenance of words. Socrates proposes that there is a connection between the way words sound and their signifiers. As an example, he gives the Greek letter ρ (rho), which for him expresses motion. A number of (Greek) words containing ρ are brought up in support of this hypothesis, for which Hermogenes provides afterwards a plethora of counter-examples.

The idea that sounds carry meaning has reappeared throughout history. Locke's *An Essay on Human Understanding* (1690) counters this idea. Leibniz's book *New Essays on Human Understanding* (1765) critiques Locke's essay. Leibniz proposes a moderate view, in which words and their referents are neither related by perfect correspondence, nor by complete arbitrariness. A detailed history of phonosemantics is presented by Genette [8], and a historical review plus recent research and developments are presented by Magnus [16].

An interesting view on the relation between sound and meaning, and the possible connection between the two, is proposed by Jakobson [11]. In Lecture VI he says: "The intimacy of connection between the sounds and the meaning of a word gives rise to the desire of speakers to add an internal relation to the external relation, resemblance to contiguity, to complement the signified by a rudimentary image". In other words, the resemblance between sound and meaning is in the ear and mind of the beholder. This may lead to a "natural selection" of words, based on the way they sound, as suggested by Otto Jespersen: "There is no denying that there are words which we feel instinctively to be adequate to express the ideas they stand for. ... Sound symbolism, we may say, makes some words more fit to survive." [12]. Firth [7] and Sapir [20] also share such a middle-ground view of sound symbolism. In their view, speech sounds carry meaning, but rather than being inherent to them, it is a result of what Firth called "phonetic habit", "an attunement of the nervous system".

3 Signifier and signified

We set out to investigate the connection between the signifier, or word form, and signified, or meaning, for English words that express emotions. Because we propose that words expressing emotions are mellifluous words, we do not seek a relation between form and exact meaning, but rather form and some aspect of the meaning - in our case, the emotion conveyed.

The signifier The signifier, in our case, can have both a written and a spoken form. A TREE is called */trE/* and written *tree* in English. The pronunciation is a sequence of sounds (phonemes). According to research in speech analysis, phonemes are not the smallest units of speech. Individual phonemes can be represented through values of a set of parameters, or features, that capture the configuration of the vocal tract that produces each sound and other acoustic features. We investigate each of these three variants of representing a word form.

letters : In English words are not pronounced as they are written. However, the way words are spelled may be closer to the words' etymological roots than their pronunciation is. As an example, the word *delight*, comes from the Old French word *delit, delitier* which in turn comes from the Latin *delectare*[3]. The letter *e* in *delight* is pronounced */i/* as in *bit*, while in its etymological roots, it is pronounced */e/* as in *bet*. Since texts are more readily available than word pronunciations, this type of word form is also the easiest to analyze.

Phoneme	Example	Transcription
AA	alarm	AH0 L AA1 R M
AE	amorous	AE1 M ER0 AH0 S
CH	charm	CH AA1 R M
EH	enchant	EH0 N CH AE1 N T
T	tickle	T IH1 K AH0 L
Y	euphoria	Y UW0 F AO1 R IY0 AH0

Table 1: *A sample of phonemes, words and their phonetic transcription*

pronunciation : Pronunciation of letters in English, especially vowels, depends on their context. Dictionaries provide a transcription of words into their phonetic equivalent. In this representation, each sound (which may correspond to one or more of a word's letters) is represented by a special symbol. We use CMU's pronunciation dictionary developed at the Carnegie Mellon University [4], which contains approximately 125,000 words and their transcriptions. The transcriptions' "alphabet" consists of 39 phonemes, and three extra digits for stress information (0 - no stress, 1 - primary stress, 2 - secondary stress). A sample of phonemes and word pronunciations are presented in Table 1.

phonetic-features : The phonemes can also be further described in terms of *phonological features* – "configurations" of the vocal tract and acoustic characteristics. From the existing phonological feature systems – [13], [9], [3] – we use the Sound Pattern of English (SPE) [3].

SPE consists of 14 binary features, which describe the tongue body position (high, back, low), tongue tip position (anterior, coronal), lips' configuration (round), configurations affecting the air flow – by constriction, vibration of vocal folds or blocking with the tongue or lips (tensed, voiced, continuant, nasal, strident) and acoustic characteristics (vocalic, consonant, silence). Examples of phonemes (also called phones) with their SPE representation are shown in Table 2.

		voc	cons	high	back	low	ant	cor	roun	tens	voic	cont	nasa	stri	sile
ae	(b*a*t)	+	-	-	+	+	-	-	-	+	+	+	-	-	-
b	(*b*ee)	-	+	-	-	-	+	-	-	-	+	-	-	-	-
iy	(b*ee*t)	-	-	+	-	-	-	-	-	+	+	+	-	-	-
m	(*mom*)	-	+	-	-	-	+	-	-	-	+	-	+	-	-
ow	(b*oa*t)	+	-	-	+	-	-	-	+	+	+	+	-	-	-
sh	(*sh*e)	-	+	+	-	-	-	+	-	-	-	+	-	+	-

Table 2: *Examples of sound representation using the SPE system*

[3] From the Online Etymology Dictionary: http://www.etymonline.com.

[4] The CMU pronunciation dictionary is freely available at http://www.speech.cs.cmu.edu/cgi-bin/cmudict . We have used version 0.6d.

The signified The signified component of our data comes from emotion tags, from two sets – a finer grained set consisting of 6 emotions, and a set consisting of 2 coarse emotion classes. Psychological research proposes the following basic emotions: {*anger, disgust, fear, joy, sadness, surprise*} [6]. We study whether analysis of word form allows us to predict whether the word expresses one of these basic emotions. Because much research in the domain of sentiment analysis works at a coarser level of emotions – *positive* and *negative* – we also study the relation between word forms and these broader emotion categories.

Resource	class	count	filtered	
WordNet Affect	anger	240	101	(25%)
	disgust	48	17	(4.2%)
	fear	134	49	(12.13%)
	joy	364	152	(37.63%)
	sadness	187	59	(14.6%)
	surprise	70	26	(6.44%)
	total	1043	404	(100%)
LIWC	negative	345	283	(59.21%)
	positive	265	195	(40.79%)
	total	610	478	(100%)

Table 3: *Word-sentiment counts from WordNet Affect and LIWC*

4 Emotion-tagged words

Assigning an emotion tag to words is not an easy task. Potentially, for any word one may perceive an emotional dimension, either directly from the word's meaning, or through the word's associations with emotionally charged words or situations.

The words we are most interested in are words that express an emotion, such as *happy, joy*. We focus on WordNet Affect [22] and LIWC [19] data because they contain words that express emotions, rather than having a semantic orientation. The word *knowledge* for example, does not have an emotion tag in WordNet Affect, but it has a positive tag in the General Inquirer data. Other resources include the General Inquirer data [5] and the list of postive and negative adjectives used by Hatzivassiloglou and McKeown [10] [6].

WordNet Affect WordNet-Affect is an extension of WordNet with affective tags. Words that have an EMOTION tag, were recently more fine-grained reannotated with one of: { *joy, fear, anger, sadness, disgust, surprise* } [22]. The choice for the six emotions comes from psychological research into human (non-verbally expressed) emotions [6].

Linguistic Inquiry and Word Count Linguistic Inquiry and Word Count (LIWC) focuses on the analysis of text and computing statistics along 82 dimensions, such as "present", "future", "space", "motion", "occupation", "physical", "metaphysical", "body "[19], based on a large dictionary that lists words under each of these dimensions. We use the words listed under "positive emotions" and "negative emotions".

Table 3 contains information about the number of unique words for each of the WordNet affect and LIWC emotions. Column 3 shows the word count for each emotion, and column 4 shows the word count after filtering morphologically related words and after verifying that the word has an entry in the CMU dictionary. The experiments are run only using words that have a pronunciation in the dictionary, to allow for comparison of performance for the different representations. We filtered morphologically related words by performing (i) stemming (using Porter's Stemmer), (ii) an extra step of cutting off suffixes (such as *-fully, -ful, -some, -ness*) to catch words with multiple suffixes, and finally (iii) word matching. We also eliminate words with the suffix "less", because the emotion the word stem expresses and the emotion expressed by the full word are different. Words with negative prefixes (un-, in-) are kept, because it is harder to detect whether a starting sequence *un* or *in* is actually a prefix or not. Also, the bigram representation, discussed below, will cover these prefixes (as opposed to the suffix *-less* for which a 4-gram representation would be necessary).

[5] The General Inquirer lexicon is freely available for research purposes from http://www.wjh.harvard.edu/ inquirer/.

[6] These and other sentiment annotated resources are available from Janyce Wiebe's web site http://www.cs.pitt.edu/˜wiebe .

5 Learning experiments

The hypothesis we explore is that word forms expressing the same emotion share sound/pronunciation characteristics – in other words, they sound similar in certain ways. The similarities may be at the smallest level – letter, sound, sound feature – or at a more complex level – letter or sound sequences, combinations of sound features. We build data representations at these three levels, and test the hypothesis using decision tree (J48, ADTree[7]) and memory based (IBK) algorithms in Weka [24], in 10-fold cross-validation experiments.

Data representation Following these considerations, we have produced a series of representations for the data, which vary along two dimensions: analyzed unit (unigrams and bigrams) and unit representation (letter, pronunciation and sound features).

We split each word into three segments – beginning segment (consisting of the first unit), ending segment (the last unit), and the middle segment which contains everything in-between. Each word is represented in terms of features for each of these three segments. For each segment, the features represent aggregated statistics for the units in this segment. For letter feature *a* in the middle segment, for example, the value is the number of occurrences of *a* in the middle segment.

An example: if we consider the word *admire* with a bigram letter representation, it will have the following segments: beginning – *ad*, middle – *dmir*, end – *re*. In its feature vector, the following features will have non-zero values: for the beginning segment – *ad*, for the middle segment – *dm, mi, ir*, for the end segment – *re*.

Table 4 shows the number of features for each data set generated for the 6 possible variations. In the table, and in the discussion that follows, we will use the abbreviations: *data sets*: WordNet Affect (W), LIWC

[7] We use Weka's MultiClassClassifier to perform multi-class classification with ADTree.

Repres.	# of Features	Repres.	# of Features
W-1Let	71	W-2Let	498
W-1P	126	W-2P	711
W-1C	42	W-2C	588
L-1Let	68	L-2Let	498
L-1P	123	L-2P	731
L-1C	42	W-2C	588

Table 4: *Number of features used in each representation method*

(L); *units*: unigram (1), bigram (2); *unit representation*: letters (Let), pronunciation/phonetic (P), SPE codes (C) levels.

For letter- and pronunciation/phonetic-based representation, the features are determined by the n-gram letter and phoneme sequences that actually appear in our list of words. For phonological features we consider the set of 14 SPE features for each word segment (beginning, middle, end). All features are numeric and their value is the number of occurrence of the feature (e.g. letter *a* or phoneme IY) in the corresponding segment of the word. The phonetic-based representation contains two extra features – for primary and secondary stress, as indicated in the pronunciation dictionary. These two features take as value the phoneme that was stressed (always corresponding to a vowel).

Results for WordNet Affect data We evaluate the quality of classification by computing the average accuracy (Acc) and average precision (P), recall (R) and F1 score (F) for each emotion class in 10-fold cross validation experiments. Our experiments have shown that dropping the features for the end of word segment has a positive impact on performance. The increase in performance may be due partly to the reduction (by approximately 33%) of the number of features.

The best results, in terms of accuracy, for the WordNet Affect data were 39.85% obtained with IB1-1Let[8] and 38.11% with IB1-1C, on word representations based only on the beginning and middle segment. In this 6-class learning problem the baseline accuracy is 37.63%, corresponding to classifying everything as *joy*, the majority class (this baseline maximizes accuracy).

Table 5 shows the best results in terms of F-score for each class (emotion) in the WordNet Affect data in the multi-class learning setting. For detailed results on each emotion class we use a baseline which guesses the class with a distribution that matches the one in the data set (this baseline balances precision and recall). The baseline F-score values are given by the distribution presented in Table 3, repeated here on row 2.

Method	anger	disgust	fear	joy	sadness	surprise
baseline	25%	4.2%	12.13%	37.63%	14.6%	6.44%
IB1-1Let	40.9%	33.3%	33.3%	49.3%	33%	9.8%
IB1-1C	42.2%	16.2%	29.5%	45.1%	31.5%	26.7%
highest	44.8%	36.8%	34.5%	55%	33%	26.7%
values	IB2-1C	IB2-1Let	IB2-1C	IB2-1C	IB1-1Let	IB1-1C

Table 5: *F1 score results on 6-class classification into WordNet Affect emotions*

The best recognized emotion from WordNet Affect's emotion classes was *joy*. Despite variation in P, R, and F values for different representations and learning algorithms, *joy* was consistently the best classified emotion. Part of this may be due to the fact that it had the most examples (37.63%). The results show statistically significant improvement over the baseline at 95% confidence level with Weka's t-test.

[8] IBK, K=1, unigram letter-based representation, following the same notation convention as in Table 4.

Results on LIWC data A selection of the best results (in terms of F1 score) for the LIWC data are presented in Table 5. The baseline F1 score is equal to the distribution of the classes, as presented in Table 3. The performance increase over the baseline is statistically significant at 95% confidence level (with Weka's t-test). For this binary classification experiment, the baseline accuracy is 56.59%, corresponding to classifying everything as *negative*, the majority class. The best results, in terms of accuracy for the LIWC data are 62.3% (IB55-1C) and 61.5% (J48-2Let).

Method	positive	negative
baseline	40.79%	59.21%
IB1-2Let	45.5%	68.2%
IB1-1P	44.1%	67.9%
highest	45.5%	74.9%
values	IB1-2Let	IB55-1C

Table 6: *F1 score results on binary classification on LIWC*

For the LIWC data, we obtained better prediction performance for words conveying a negative emotion. There are also more words expressing negative emotions in our data set.

IBK, which classifies a word based on its similarity with neighbouring words, outperforms other classifiers in finding the best results for both the binary and the 6-class learning problems. This supports the idea that words expressing the same emotion have more in common with each other than with words expressing other emotions.

6 The sounds of emotions

Happy words sound more like other happy words than like words expressing other emotions. But do they really sound happy?

In order to verify whether such features are indeed perceived as expressing the emotion we consider, we look into research on recognizing emotions in human speech. The type of data used in such work are recordings of (usually, multi-word) utterances, whose sound signal is represented through a variety of features (such as pitch, energy, tone contour) [21],[4], [18]. Lee et al. [14] introduce five broad phoneme classes – vowel, stop, glide, nasal, fricative – to help in classifying utterances into 4 classes – angry, happy, neutral and other. In learning experiments using Hidden Markov Models, they note that using phoneme classes in addition to the more traditional signal features leads to better emotion recognition. In particular, vowel sounds are good emotion indicators, and furthermore different vowels have different effects, possibly because of articulatory constraints: "less constricted low vowels such as /AA/ show greater effects than do high vowels like /IY/". There are no details as to which vowels are predictive of which emotion class, but it is not just the presence or absence of a vowel that is useful for predicting the class, but also prosodic features related to its pronunciation [15].

Whissell [23] analyzed phonologically transcribed text samples from song lyrics, poetry, word lists and advertisements) for correlations between phonemes and language emotionality. Phonemes were grouped into 8 classes, based on two dimensions – Pleasantness and Activation. Support for this grouping was given through experiments using phonemes as part of non-words. Here are the classes and a sample of their assigned phonemes: Pleasantness – /AY/-*high*, /DH/-*this*; Cheeriness – /AA/-*father*, /AY/-*high*, /CH/-*chip*, /F/, /V/; Softness – /TH/-*thumb*, /EH/-*bet*, /L/, /M/; Activation – /AA/-*far*, /OY/-*voice*; Nastiness – /ER/-*her*, /UW/-*cool*, /NG/; Unpleasantness – /AW/-*cow*, /OW/-*bone*; Sadness – /AW/-*cow*, /B/; Passivity – /AE/-*hat*, /K/, /L/ [9].

Let us now look in a bit more detail at some of the most salient features in our data representation, as identified by the tree-based algorithms. Negative words in LIWC data are characterized by vocalic beginning and phonemes pronounced with the tongue body in back position(e.g. /CH/, /NG/, /G/, /AA/, /AH/), as in *angry*. Such phonemes appear in the Nastiness category [23]. Positive words by starting phonemes pronounced with the tongue body in high and the tip not coronal position (e.g. /IY/, /K/, /P/) and at most two phonemes pronounced with the tongue in back position in the middle segment (e.g. *improve, kind*). /IY/ appears in the Softness, Pleasantness and Cheerful category category, but /K/ and /P/ appear in the Unpleasantness and Passive ones. Happy words in the WordNet Affect data start with phonemes which are not continuant and the tongue tip is not in anterior position (e.g. /CH/, /NG/, /K/) and the body contains tensed phonemes (e.g. /AA/, /AW/, /EY/) (e.g. *charming*. Words expressing sadness start with non-consonantal phonemes pronounced with the tongue body in back position (e.g. /AW/, /OW/, /UH/).

We observe parallels between the features found most discriminating by the decision tree algorithms, and the phonemes previously established in the literature as having emotional connotations. We also note that it is the effect of several phonemes that gives a word its "emotion" sound. In future work we will determine a representation that captures best the interactions and relations between sounds in a word.

7 Conclusion

We have investigated the properties of word-forms, to learn whether we can automatically predict the emotion a word expresses based on various representations of its form. The results show that all the representations used – word spelling, pronunciation, phonetic features – are useful for determining that words expressing the same emotion are alike in certain ways.

These results answer half of the question we had set out to investigate – whether words sound like the emotion conveyed. The other half is whether what happy words have in common is what makes them sound happy. The answer to this question is harder to find, because of subjectivity of perception and bias from the meaning component. We have found interesting parallels with features used in classifying emotion words and emotional sound characteristics found in related work. Future work on larger text segments annotated with emotion and future developments in emotion recognition in speech analysis will help provide a more rigorous answer to this part of the question.

We plan to experiment with alternative word representations – such as syllables, which are considered the phonological "atoms" of words – and to determine which part of the word is most expressive from the point of view of the emotion conveyed. Research based on the words' etymological roots may show us if the link between form and meaning gets stronger as we go back in time. Next step is to expand the study to languages other than English, and to longer text units, such as blogs. In speech emotions are detectable, and the speaker conveys these through tone and other prosodic features. It would be interesting to see whether we can identify "sub-word" level features useful for detecting emotions in blogs.

Acknowledgments Partial support for this work is given by RALI (Guy Lapalme et al.), Université de Montreal. We thank the three anonymous reviewers for their very good, insightful comments.

References

[1] S. Bedgley. Strawberry is no blackberry: Building brands using sound, 2002. http://online.wsj.com/article/0,,SB1030310730179474675.djm,00.html.

[2] H. Bredin. Onomatopoeia as a figure and a linguistic principle. *New Literary History*, 27(3):555–569, 1996.

[3] N. Chomsky and M. Halle. *The Sound Pattern of English*. MIT Press, 1968.

[4] R. Cowie, E. Douglas-Cowie, N. Tsapatsoulis, G. Votsis, S. Kolias, W. Fellenz, and J. Taylor. Emotion recognition in human-computer interaction. *IEEE Signal Processing Magazine*, 18(1):32–80, 2001.

[5] F. de Saussure. *Cours de linguistique générale*. Harrassowitz, Wiesbaden, 1916.

[6] P. Ekman. An argument for basic emotions. *Cognition and Emotion*, 6:169–200, 1992.

[7] J. R. Firth. Modes and meaning. In *Papers in linguistics 1934-1951*. Oxford University Press, London, 1951.

[8] G. Genette. *Mimologiques: voyage en Cratylie*. Seuil, Paris, 1976.

[9] J. Harris. *English Sound Structure*. Blackwell, 1994.

[10] V. Hatzivassiloglou and K. McKeown. Predicting the semantic orientation of adjectives. In *Proceedings of the 35th ACL/8th EACL*, pages 174–181, Madrid, Spain, 1997.

[11] R. Jakobson. *Lectures on Sound and Meaning*. MIT Press, Cambridge, MA, 1937.

[12] O. Jespersen. *Language - its Nature, Development and Origin*. George Allen & Unwim Ltd., London, 1922.

[13] J. King and P. Taylor. Detection of phonological features in continuous speech using neural networks. *Computer speech and language*, 14(4):333–353, 2000.

[14] C. M. Lee, S. Yildirim, M. Bulut, A. Kazemzadeh, C. Busso, Z. Deng, S. Lee, and S. Narayanan. Emotion recognition based on phoneme classes. In *Proc. of ICSLP*, 2004.

[15] L. Leinonen, T. Hiltunen, I. Linnankoski, and M.-L. Laakso. Expression of emotional-motivational connotations with a one-word utterance. *Acoustical Society of America Journal*, 102:1853–1863, 1997.

[16] M. Magnus. *What's in a Word? Studies in Phonosemantics*. PhD thesis, University of Trondheim, Trondheim, Norway, 2001.

[17] M. Nänny and O. Fischer. Iconicity: Literary texts. In K. Brown, editor, *Encyclopedia of Language and Linguistics*, volume 5, pages 462–472. Elsevier, Oxford, 2nd edition, 2006.

[18] A. Nogueiras, A. Moreno, A. Bonafonte, and J. B. M. no. Speech emotion recognition using Hidden Markov Models. In *Proc. of Eurospeech*, 2001.

[19] J. W. Pennebaker, R. J. Booth, and M. E. Francis. Linguistic Inquiry and Word Count: LIWC. [Computer software]. Mahwah, NJ: Erlbaum., 2001.

[20] E. Sapir. A study in phonetic symbolism. *Journal of Experimental Psychology*, (12):225–239, 1929.

[21] K. R. Scherer. Vocal communication of emotion: A review of research paradigms. *Speech Communication*, 40(1):227–256, 2003.

[22] C. Strapparava, A. Valitutti, and O. Stock. The affective weight of the lexicon. In *Proceedings of the 5th International Conference on Language Resources and Evaluation (LREC 2006), Genoa, Italy*, 2006.

[23] C. Whissell. Phonoemotional profiling: a description of the emotional flavour of english texts on the basis of the phonemes employed in them. *Perceptual and Motor Skills*, 91(2):617–648, 2000.

[24] I. H. Witten and E. Frank. *Data Mining: Practical machine learning tools and techniques*. Morgan Kaufmann, San Francisco, 2nd edition edition, 2005.

[9] Table 1 at http://www.trismegistos.com/IconicityInLanguage/Articles/WhisselPlath/index.html